A Population Proxy Index (PPI) for the Selection of Habitat Reserves

Bala Krishnamoorthy

joint work with
Natalie Baerlocher, David Allen
NSF-UBM Project

INFORMS 2009 San Diego
The Habitat Selection Problem

• Habitat reserve design for a territorial disperser
e.g., the Northern Spotted Owl (NSO)

• Binary Habitat
  Green = Suitable
  Beige = Unsuitable

• Questions:
  – Which habitat areas will best preserve the species?
  – Which habitat areas will best augment an existing reserve area?
A Sample Habitat Landscape
Life Cycle Parameters of the NSO

- Home territory ≈ 900 hectares
- Adult survival rate ≈ 0.91
- Birth rate and juvenile survival rate combined ≈ 0.20
- Juvenile dispersal process - sensitive to fragmentation
- Mean dispersal distance: 15km
Previous Work

- On various subspecies of spotted owls
  - Lamberson et al. (1992, 1994) – patch size and spacing
  - Bond et al. (2004) – California spotted owls
  - Stauffer et al. (2005) – Bayesian methods
  - McDonald et al. (2006) – discrete choices

- Moilanen (2005) – non-linear distribution models

- IP: several recent studies on multiple species
Dynamic Population Model

- Model habitat as array of cells, each cell the size of a home territory
- Habitat cells are either suitable or unsuitable
  \[ H_{ij} = 1 \text{ or } 0 \]
- Randomly seed adults (nesting pairs)
- Model several cycles of mortality, birth, dispersal, etc., until population reaches equilibrium
- Compute average occupancy rate for each cell as a measure of the "value" of the cell
Limitations of Dynamic Model

- Captures species life history dynamics
- Not useful for constrained optimization
- Need an easy to evaluate function that measures the “suitability” of a habitat for the survival of the species
Index Function Modeling Problem

- Develop an explicit landscape function (PPI) that measures habitat “values” similar to the dynamic population model (i.e., calculate occupancy rates)

- Use the explicit function to formulate constrained optimization problems in resource management
The Population Proxy Index (PPI)

- Estimate of the long-term population of a cell
- Calculated as a “colonization probability”
- Each suitable cell has an independent probability of “colonizing” the focal cell
- The probability of colonization is a function of distance
Colonization

• If the location of the natal cell is \((k, l)\) then the probability that a dispersal attempts to claim cell \((k + i, l + j)\) is \(p_{ij}\)
  
  – depends on distance \(d_{ij}\)
  – but is independent of \(k, l\)

• The probability that at least one neighboring cell manages to colonize cell \((i, j)\) is given by

\[
1 - \prod_{k,l} (1 - p_{kl} H_{i-k,j-l})
\]
Population Proxy Index

- **PPI** for cell \((i, j)\) is the product of the habitat value for that cell, \(H_{ij}\), and the probability that at least one (other) cell will colonize \((i, j)\):

\[
PPI_{ij} = H_{ij} \left[ 1 - \prod_{k,l} (1 - p_{kl} H_{i-k,j-l}) \right]
\]

- Recall \(H_{i-k,j-l} = 0\) or 1,

\[
\ln(1 - p_{kl} H_{i-k,j-l}) = H_{i-k,j-l} \ln(1 - p_{kl})
\]

- Let \(L_{kl} = \ln(1 - p_{kl})\) (log-dispersal propensity)
Population Proxy Index

\[ PPI_{ij} = H_{ij} \left( 1 - e^{\sum_k \sum_l (H_{i-k,j-l} - L_{kl})} \right) \]

- \( L_{kl} \) is obtained through a *tuning* process
  - use regression or linear programming to fit real data
  - we use data from simulations

- For the entire landscape, \( PPI = \sum_{all\ i,j} PPI_{ij} \)

- Highly nonlinear
Example Calculation

\[ L = \]

\[ PPI_{ij} = 1 \times (1 - e^{-0.3-0.5-0.5-0.3}) \approx 0.8 \]
The Motivating Problem

- Want to determine which areas of the habitat to maintain, and which ones could be cleared or utilized for other activities (e.g., logging)
- the species should survive in the altered habitat
- Develop a constrained optimization problem

\[
\begin{align*}
\max & \quad PPI(x_1, x_2, \ldots, x_n) \\
\text{s.t.} & \quad c_1x_1 + c_2x_2 + \ldots + c_nx_n \leq G \\
& \quad x_j \in \{0, 1\} \text{ for } j = 1, \ldots, n
\end{align*}
\]

- e.g., can maintain as green at most 15 candidate squares

\[\Rightarrow \quad c_1, \ldots, c_n = 1, \quad G = 15\]
Quadratic Model

\[ PPI = \sum_i \sum_j H_{ij} \left[ 1 - \exp \left( \sum_k \sum_l H_{i-k,j-l} L_{kl} \right) \right] \]

- need to use (nicer) approximations of \( PPI_{ij} = H_{ij}(1 - e^{t_{ij}}) \)

where \( t_{ij} = \sum_k \sum_l (H_{i-k,j-l} L_{kl}) \)

- notice that \( e^t \approx 1 + t \) for small \( t \), so define

\[ QPPI_{ij} = -H_{ij} \sum_k \sum_l (H_{i-k,j-l} L_{kl}) \]

- \( QPPI = \sum_{i,j} QPPI_{ij} \) is positive semidefinite, so easier to solve the optimization problem
Example Calculation – Quadratic Model

\[ L = \begin{bmatrix}
-0.3 & -0.5 & -0.3 \\
-0.5 & 0 & -0.5 \\
-0.3 & -0.5 & -0.3
\end{bmatrix} \]

\[ QPPI_{ij} = -1 \times (-0.3 - 0.5 - 0.5 - 0.3) = 1.6 \]
Graph of $PPI_{ij}$ vs $QPPI_{ij}$

- $QPPI_{ij}$ is not a good approximation of $PPI_{ij}$
- note that $1 - e^t \leq -t$ and also $1 - e^t \leq 1$
Linearized Model

- use several linear pieces to represent $1 - e^{t_{ij}}$
Linearized Model

• Let $y_{ij}$ represent $1 - e^{t_{ij}}$. To use $r$ pieces, we write

$$y_{ij} \leq P_k(t_{ij}), \ k = 1, \ldots, r,$$

where

$$P_1(t_{ij}) = 1$$

$$P_2(t_{ij}) = -t_{ij}$$

$$P_3(t_{ij}) = -0.5 \cdot t_{ij} + 0.35 \ (say)$$

• define $PLQPPPI_{ij} = H_{ij} \times y_{ij}$, and add these constraints

• need to further linearize quadratic terms $(x_{ij} \times y_{ij})$ ...
Fully Linearized Model

- let $z_{ij}$ represent the quadratic term $x_{ij} \times y_{ij}$
- define $PLPPI_{ij}$ by replacing $x_{ij} \times y_{ij}$ with $z_{ij}$ in $PLQPPI_{ij}$
- add the constraints

\[
\begin{align*}
  z_{ij} & \leq y_{ij} + (1 - x_{ij}) \\
  z_{ij} & \leq x_{ij}
\end{align*}
\]

recall: $x_{ij}$ is 0 or 1

- $PLPPI = \sum_{i,j} PLPPI_{ij}$
Computation

• 125 × 75 landscape map, 51 × 51 dispersal range

• Vary # candidate habitat cells \((n)\) and max # of candidate cells that can be kept green \((G)\)

\[
\begin{align*}
\text{max} & \quad PPI(x_1, x_2, \ldots, x_n) \\
\text{s.t.} & \quad x_1 + x_2 + \ldots + x_n \leq G \\
& \quad x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n
\end{align*}
\]

with \(QPPI\) and \(PLPPI\) replacing \(PPI\)

• Solve IP using CPLEX
Candidate Habitats

- tried many different configurations of candidate habitats (red cells)
Quadratic Model – varying $n$

Figure 1: Run times (in hours) of the quadratic model for varying total # cells in the two (red) patches, with $G = 25$. 
Figure 2: Run times (in hours) of the quadratic model for varying maximum # candidate cells allowed to be green, with $n = 80$. 
How many pieces are needed?

- evaluate $PPI$ at optimal solutions obtained using $PLPPI$ with various numbers of linear pieces

- 3 Pieces are sufficient
Figure 3: # BB nodes for solving the linearized model for varying total # cells in the two (red) patches, with $G = 25$.

Note: # constraints is large for large $n$
Linearized Model (3 Pieces) – varying $G$

Figure 4: # BB nodes for solving the linearized model for varying maximum # candidate cells allowed to be green, with $n = 80$. 
Applicability of PPI Model

- Natural to observe decreasing probability with distance
- Geometry of landscape can be arbitrary (not necessarily be square patches)
- Need info of whether the cell can colonize the focal cell
- Can include various size and geometry restrictions as constraints in the linearized IP
Further Work

- compare with real data

- Graded habitats – $0 \leq H_{ij} \leq 1$, with higher value indicating a more suitable cell

  \[ \ln(1 - p_{kl}H_{ij}) \approx H_{ij} \ln(1 - p_{kl}) \]

- Maintain PPI value (above $M$), while minimizing the number of candidate cells that have to be kept green:

  \[ \min x_1 + x_2 + \ldots + x_n \]

  \[ \text{s.t.} \quad PPI(x_1, x_2, \ldots, x_n) \geq M \]

  \[ x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n \]

- Model multiple, and competing species